区間代数用データと誤差消失 (数学的ソフトウェアの評価)

AUTHOR(S):

平野, 菅保

CITATION:

平野, 菅保. 区間代数用データと誤差消失 (数学的ソフトウェアの評価). 数理解析研究所講究録 1979, 359: 100-124

ISSUE DATE:

1979-07

URL:

http://hdl.handle.net/2433/104510

数理解析研究所講究録
359巻 1979年 100-124

区間代数用データと誤差消失

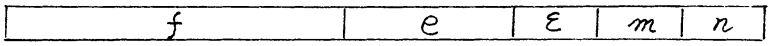
NTIS　　平野菅保

0 序

実際に計算をするときの数値は有限桁である。したがって真の実数で計算しているわけではない。そこで、計算をするときの数値には「はば」すなわち「誤差」,「無効桁」を考えねばならず、それらの数値を用いて計算すると、実際には計算されている桁でも無効桁を考えねばならなくなる。また真の「はば」も実数であることを考えると、厳密な意味で計算結果の数値の無効桁を表現することは困難である。そこで次に述べるような簡単なデータ構造と演算を考えてみる。

この方法により計算結果の有効桁を評価すると、真の誤差より桁違いに過大に誤差を評価する場合がしばしばある。これは計算途中に起こる誤差の消失に起因することであり、計算結果の誤差を得ようとするときには、このことを特に注意しなければならない。

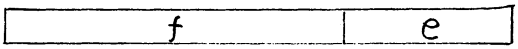
1　データの構造

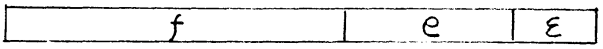


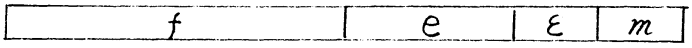
f：仮数部（符号を含む。）　　m：εを決定したデータ番号

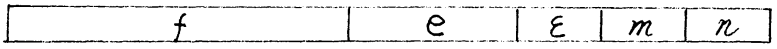
e：指数部（符号を含む。）　　n：演算回数

ε：fの誤差による無効桁数

Ⅰタイプ　 

Ⅱタイプ　 

Ⅲタイプ　 

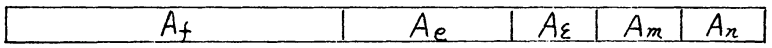
Ⅳタイプ　 

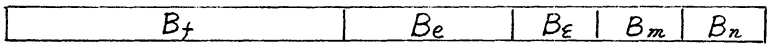
Ⅰタイプが一般の型であるが、およその無効桁を知るためにはⅡタイプを用い、その無効桁が入力、又は計算途中のどの数値によるものかを知るためにⅢタイプを用いる。又どの程度演算された数値かを知るためにはⅣタイプを用いる。

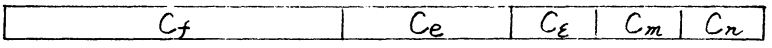
2　演算

f，e，ε，m，nは整数のように読めたり、書き込んだりできること。

C＝A⊕B　　⊕：四則演算の記号

A　 

B　 

C　 

C_f，C_eのデータは一般に行なわれている演算と同様に求め

る。又丸めは4捨5入とする.

加減算 $C_\varepsilon = max\,[\,A_\varepsilon + A_e,\ B_\varepsilon + B_e\,] - C_e$

$C_m = A_m \qquad A_\varepsilon + A_e \geqq B_\varepsilon + B_e$

$C_m = B_m \qquad A_\varepsilon + A_e < B_\varepsilon + B_e$

$C_n = max\,(A_n, B_n) + 1$

乗除算 $C_\varepsilon = max\,(A_\varepsilon, B_\varepsilon) + (A_e + B_e) - C_e$ （乗算）

$C_\varepsilon = max\,(A_\varepsilon, B_\varepsilon) + (A_e - B_e) - C_e$ （除算）

$C_m = A_m \qquad A_\varepsilon \geqq B_\varepsilon$ 注：$C_\varepsilon = -1$ のときは

$C_m = B_m \qquad A_\varepsilon < B_\varepsilon$ $C_\varepsilon = 0$ とする。

$C_n = max\,(A_n, B_n) + 1$

例1 IIタイプ ——：無効桁

加算

C_f	C_e	C_ε	A_f	A_e	A_ε	B_f	B_e	B_ε
+1.6595	+2	4	+2.1100	-1	4	+1.6574	+2	4
$+1.\underline{6595}\times10^{+2}$			$+2.\underline{1100}\times10^{-1}$			$+1.\underline{6574}\times10^{+2}$		

$C_\varepsilon = max\,(A_\varepsilon + A_e, B_\varepsilon + B_e) - C_e = max\,(3, 6) - 2 = 4$

減算

C_f	C_e	C_ε	A_f	A_e	A_ε	B_f	B_e	B_ε
+2.1100	-1	4	+3.7452	+1	2	+3.7241	+1	2
$+2.\underline{1100}\times10^{-1}$			$+3.74\underline{52}\times10^{+1}$			$+3.72\underline{41}\times10^{+1}$		

$C_\varepsilon = max\,(A_\varepsilon + A_e, B_\varepsilon + B_e) - C_e = max\,(3, 3) - (-1) = 4$

乗算

$$\begin{bmatrix} C_f & C_e & C_\varepsilon \end{bmatrix} \begin{bmatrix} A_f & A_e & A_\varepsilon \end{bmatrix} \begin{bmatrix} B_f & B_e & B_\varepsilon \end{bmatrix}$$

$$\begin{bmatrix} +1.5508 & +2 & 2 \end{bmatrix} = \begin{bmatrix} +1.2453 & +1 & 2 \end{bmatrix} \times \begin{bmatrix} +1.2453 & +1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} +1.55\underline{08} \times 10^{+2} \end{bmatrix} = \begin{bmatrix} +1.24\underline{53} \times 10^{+1} \end{bmatrix} \times \begin{bmatrix} +1.24\underline{53} \times 10^{+1} \end{bmatrix}$$

$$C_\varepsilon = \max(A_\varepsilon, B_\varepsilon) + (A_e + B_e) - C_e$$

$$= \max(2, 2) + (1+1) - 2 = 2$$

除算

$$\begin{bmatrix} C_f & C_e & C_\varepsilon \end{bmatrix} \begin{bmatrix} A_f & A_e & A_\varepsilon \end{bmatrix} \begin{bmatrix} B_f & B_e & B_\varepsilon \end{bmatrix}$$

$$\begin{bmatrix} +2.0520 & +1 & 3 \end{bmatrix} = \begin{bmatrix} +2.5324 & +1 & 2 \end{bmatrix} \div \begin{bmatrix} +1.2341 & 0 & 3 \end{bmatrix}$$

$$\begin{bmatrix} +2.05\underline{20} \times 10^{+1} \end{bmatrix} = \begin{bmatrix} +2.53\underline{24} \times 10^{+1} \end{bmatrix} \div \begin{bmatrix} +1.23\underline{41} \times 10^{0} \end{bmatrix}$$

$$C_\varepsilon = \max(A_\varepsilon, B_\varepsilon) + (A_e - B_e) - C_e$$

$$= \max(2, 3) + (1-0) - 1 = 3$$

例2　4元連立1次方程式を解いて、解 X_i ($i = 1, 2, 3, 4$) が得られ、その解を与えられた方程式に代入して、その式を満足しているか調べるときに用いた。
係数、定数ともに最後の桁は無効桁とした。

例3　3重根の解を含む3次方程式をニュートン法で求める途中の近似解2つを、与えられた方程式に代入した。この2つの近似解は2～3桁の精度しかないが、関数値は無効桁のみになっている。又係数の最後の桁は無効桁とした。

3　データの長さ（ビット数）

fがSingleでeがQuadrupleの場合，εの値、すなわち無効桁がfの桁数より大きい場合等がある。非線形の方程式を解くには、仮数部fはSingleでも、指数部eの範囲が大きくないと解法を考えるとき制約をうける。

	f	e	ε	m	n	指数部の範囲	10進桁数
Single	32	8	8	8	8	$10^{\pm 38}$	9.3
Double	64	16	16	⋮	16	$10^{\pm 9864}$	18.9
Quadruple	128	32	⋮	⋮	32	$10^{\pm 6464569 94}$	38.2
Octuple	256	⋮	⋮	⋮	⋮	⋮	76.7
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

4　誤差の消失

計算に用いる数値の無効桁に比較して、計算結果の無効桁が増加するような計算で、計算に用いる数値の下位に無効桁を多数桁加えてこの演算方式で計算すると、無効桁が単に増加するだけのようであるが、誤差の従属性により誤差が消失し、桁を増加しないで計算したときと異なって、実際には無効桁が減少している。すなわちこの演算方式で計算結果の無効桁を知る方法では、例4以後に述べるように、この演算方式を使用してはならない場合があることに注意しなければならない。

例4　$f(X) = \cos X = 1 + \sum_{n=1}^{\infty} (-1)^n X^{2n} / (2n)!$

微分方程式とすれば $d^2f/dX^2=-f$, $X=0.$ のとき $df/dX=0$, $f=1.0$ の場合である。

6桁の $X=4.47214$ を用いて6桁で計算すると、計算途中で上位2桁（10と1の桁）が桁落ちして（ ）の真値に比較して下位2桁が無効桁となっている。これは正しい。しかし、6桁のXの下位に4桁任意の数値，無効桁3142を加えて10桁として同様に計算すると、——のように無効桁はならず、——の1桁目4は無効桁ではない。すなわち誤差は消失している。

例5　$f(X)=e^X$ を $X=1.0$ で数値微分する。きざみはばを 2^{-n} とし、前進と後退の1次微分の値をまず求める。（2進数36ビット，仮数部28ビットの計算機で計算した。）

相対誤差は $n=14$ でともにもっとも小さく（7.4×10^{-6}）、nが14より小さければ打ち切り誤差，14より大きければ丸めの誤差で相対誤差は大きくなる。次いで1次微分の数値を用いて2次微分の値を求めると、1次微分では相対誤差が 3.9×10^{-3} で大きかった $n=7$ で逆に誤差は小さくなり、相対誤差は 7.4×10^{-6} となっている。これは打ち切り誤差の消失が起こったのである。

例6　$\displaystyle\int_0^{\pi}\frac{1+\sin x}{2+\cos x}\,dx$

上式の数値積分を前半周期と後半周期を別々に求めると、

シンプソン公式を用いた方が台形公式を用いた場合より結果の精度がよい．すなわち無効桁が少ない．ところがそれらの和を求めると，台形公式を用いた方の精度がよくなり，無効桁は少なくなる．これは，台形公式の場合，前半周期と後半周期とでは打ち切り誤差が絶対値上位数桁一致して，符号が逆になっているために，誤差が消失しているからである．シンプソン公式の場合にも同様な傾向はでている．

例7　2次方程式を解く場合，2つの解の絶対値が極端に異なると，式　$(-b\pm\sqrt{b^2-4ac})/(2a)$　を用いて計算すると，絶対値の小さい解の無効桁が多くなる．ところが係数 a, b, c の下位の桁に任意の数値を加えて，多数桁で計算すると，絶対値の小さい解も，誤差が消失して，無効桁は増えない．計算桁数は5桁と10桁である．真の解は $\pi\times10^{2}$, $\pi\times10^{-2}$ である．

例8　3元連立1次方程式をガウスの消去法で解く。$a_{11}^{(0)}=\varepsilon$ のみが誤差のように絶対値が小さいとき，第1回目の消去でピボットとして $a_{11}^{(0)}$ を用いて単精度（10進4桁）で計算すると，それ以後のガウスの消去計算により無効桁のみの解が得られる．ところが $a_{11}^{(0)}$ で $a_{12}^{(0)}$, $a_{13}^{(0)}$, $c_{1}^{(0)}$ を除すまでの計算を単精度で行ない，それ以後任意の数値を下位の桁に加えて2倍精度（10進8桁）ですべて計算すると，得られた解の

精度は与えられた係数の精度程度ある。すなわち丸めの誤差の消失が起こる。

例9　1元3次方程式

真の解は$\pi\times10^4$, $\pi\times10^2$, πである。

一般に、方程式に含まれている解のうち、絶対値の大きい解が先に得られると、その解を用いて次数を下げる計算のときに桁落ちが起こって、次数の1つ下がった方程式の係数に誤差が大きく入る。すなわち無効桁が増加する。そこで1元3次方程式の係数(10進4桁)の下位の桁に任意の数値を加えて収束計算をすると、(解が$\pi\times10^4$の場合は10桁収束させるし、$\pi\times10^2$の場合は6桁収束させる。)次数を下げる計算で誤差も消失して、求められた1元2次方程式の係数は4桁精度があり、無効桁は増加しない。したがって次に得られる2つの解の精度は4桁あり、無効桁は増加していない。

例10　次の1元3次方程式を多項式に展開して解を求めると、計算解X_2, X_3の無効桁(この場合誤差は虚数の形になっている。)は2桁強である。又係数の精度は10進4桁であり

真の解は$X_1=e\times10$, $X_2=X_3=\pi$である。

$$(X-3.142)(X^2-9.870)=(5.507X-17.30)^2$$

ところが、多項式に展開する前の係数の下位の桁に任意の数値を加えて8桁として計算すると、誤差は消失して、無効

桁は増加していない．重根の解 $X_2 = X_3$ も4桁の精度がある。

例11　行列 $a_{ij} = min(i,j)$ $(i,j = 1, 2, \cdots\cdots, 11, 12)$ の固有値を求めるためにダニレフスキー法でコンパニオン行列を作ると、第12列目の要素がその固有値を含む1元12次方程式の係数になっている．（2進数48ビット，仮数部38ビットの計算機で計算した。）

はじめに単精度でコンパニオン行列を作り、1元12次方程式の係数を計算機から出力するときは2倍精度にし、その2倍精度の係数をもつ1元12次方程式を2倍精度で解いた。1元12次方程式の係数の精度は単精度程度求められているが、解の精度は悪い。

次に行列 $a_{ij} = min(i,j)$ の各要素の下位の桁に任意の数値を加えて2倍精度にし、2倍精度でコンパニオン行列を作り、1元12次方程式を2倍精度で解く。1元12次方程式の係数の精度は前の場合と同様に単精度程度であるが、解の精度は前の場合よりよくなっている．すなわち1元12次方程式の係数が含む誤差は互いに独立ではなく、誤差の消失が起こっている。

例12　固有値 $\lambda_{1T}, \lambda_{2T}, \lambda_{3T}$，それに対応する固有ベクトルを V_{1T}, V_{2T}, V_{3T} とし、次式により行列 A_T を作り、小数点

以下5桁目を4捨5入して行列A_0を作った。

$$A_T = \sum_{i=1}^{3} \lambda_{iT} V_{iT} V_{iT}^T \qquad V_{iT}^T : V_{iT} \text{の転置行列}$$

べき乗法を用い、10進4桁で固有値を$\lambda_{1T}, \lambda_{2T}, \lambda_{3T}$の順に求めると、$\lambda_{1T}, V_{1T}$は精度よく求められるが、$\lambda_{3T}, V_{3T}$の精度は非常に悪い。しかし行列$A_0$の各要素の下位の桁に任意の数値（ここでは0）を加えて行列$\bar{A}_0$を作り、10進10桁で計算すると、λ_{3T}の精度は悪いが、V_{3T}の精度はよい。これも誤差の消失による。

例 2　　——— : 無効桁

a_{11} +3.85000E+0	a_{12} +9.25000E-1	a_{13} 0.	a_{14} +1.00000E-2	c_1 +3.75925E+1
a_{21} +9.25000E-1	a_{22} +1.24000E+1	a_{23} +9.25000E-1	a_{24} +4.33000E-2	c_2 -3.09000E+0
a_{31} 0.	a_{32} +9.25000E-1	a_{33} +3.85000E+0	a_{34} +1.00000E-2	c_3 -3.94075E+1
a_{41} +1.00000E-2	a_{42} +4.33000E-2	a_{43} +1.00000E-2	a_{44} +1.76000E-4	c_4 -1.32700E-2

X_1 +1.00058E+1 (10.)	X_2 +1.10079E-1 (0.1)	X_3 -9.99428E+0 (-10.)	X_4 -1.03131E+2 (-100.)

( ) 内は真の解

1 式

(1)= $a_{11} X_1$ = 38.5223

(2)= $a_{12} X_2$ = 0.101823

(1)+(2)= 38.6241

(4)= $a_{14} X_4$ = -1.03131

(1)+(2)+(4)= 37.5928

$-c_1$ = -37.5925

合計 0.000300000

3 式

(2)= $a_{32} X_2$ = 0.101823

(3)= $a_{33} X_3$ = -38.4780

(2)+(3)= -38.3762

(4)= $a_{34} X_4$ = -1.03131

(2)+(3)+(4)= -39.4075

$-c_3$ = 39.4075

合計 0.

2式

(1)= $a_{21}X_1$=	9.25537
(2)= $a_{22}X_2$=	1.36498
(1)+(2)=	10.6204
(3)= $a_{23}X_3$=	-9.24471
(1)+(2)+(3)=	1.37569
(4)= $a_{24}X_4$=	-4.46557
(1)+(2)+(3)+(4)=	-3.08988
$-c_2$=	3.09000
合計	0.000120000

4式

(1)= $a_{41}X_1$=	0.100058
(2)= $a_{42}X_2$=	0.0047664 2
(1)+(2)=	0.104824
(3)= $a_{43}X_3$=	-0.0999428
(1)+(2)+(3)=	0.00488120
(4)= $a_{44}X_4$=	-0.0181511
(1)+(2)+(3)+(4)=	-0.0132699
$-c_4$=	0.0132700
合計	0.000000100000

— : 無効桁

例 3

$$f_3(X)=(X-3)^3 = X^3 - 9.0000000X^2 + 27.000000X - 27.000000 = 0$$

(a_3 a_2 a_1 a_0)

X	+2.9844569	+2.9890621
a_2	-9.0000000	-9.0000000
$X+a_2$	-6.0155431	-6.0109379
$(X+a_2)X$	-17.953129	-17.967067
a_1	+27.000000	+27.000000
$(X+a_2)X+a_1$	+9.0468710	+9.0329330
$((X+a_2)X+a_1)X$	+26.999997	+26.999998
a_0	-27.000000	-27.000000
$f_3(X)$	-0.00000300000000	-0.00000200000000

—— : 無効桁

例 4

	X=4.47214	X=4.472143142
	cosX=	cosX=
1	1.00000	1.000000000
$-X^2/2!$	-10.0000	-10.00003214
$+X^4/4!$	+16.6667	+16.66677380
$-X^6/6!$	-11.1112	-11.11121824
$+X^8/8!$	+3.96828	+3.968304983
$-X^{10}/10!$	-0.881842	-0.8818483861
$+X^{12}/12!$	+0.133613	+0.1336138213
$-X^{14}/14!$	-0.0146827	-0.01468288469
$+X^{16}/16!$	+0.00122356	+0.001223577657
$-X^{18}/18!$	-0.0000799716	-0.00007997265289
$+X^{20}/20!$	+0.00000420904	+0.000004209100522
$-X^{22}/22!$	-0.000000182210	-0.0000001822127295
$+X^{24}/24!$		+0.000000006601931708
$-X^{26}/26!$		-0.0000000002031370130
合計	-0.237984	-0.2379414082
	(-0.237944.....)	

— : 無効桁

例 5　　$f(x)=e^x$

n	$\{f(1.0+2^{-n})-f(1.0)\}/2^{-n}$		$\{f(1.0)-f(1.0-2^{-n})\}/2^{-n}$		$\{\nabla f(1.0)-\nabla f(1.0-2^{-n})\}/2^{-n}$	
1	3.5268145	3.0E-1	2.1391211	-2.1E-1	2.7753868	2.1E-1
2	3.0882444	1.4E-1	2.4051273	-1.2E-1	2.7324686	5.2E-3
3	2.8954802	6.5E-2	2.5552523	-6.0E-2	2.7218227	1.3E-3
4	2.8050256	3.2E-2	2.6350780	-3.1E-2	2.7191620	3.2E-4
5	2.7612009	1.6E-2	2.6762476	-1.5E-2	2.7185059	8.2E-5
6	2.7396297	7.9E-3	2.6971550	-7.8E-3	2.7183838	3.8E-5
7	2.7289276	3.9E-3	2.7076912	-3.9E-3	2.7182617	-7.4E-6
8	2.7235947	2.0E-3	2.7129822	-1.9E-3	2.7167969	-5.5E-4
9	2.7209320	9.7E-4	2.7156372	-9.7E-4	2.7109375	-2.7E-3
10	2.7196045	4.9E-4	2.7169495	-4.9E-4	2.7187500	1.7E-4
11	2.7189331	2.4E-4	2.7176514	-2.3E-4	2.6250000	-3.4E-2
12	2.7186279	1.3E-4	2.7180176	-9.7E-5	2.5000000	-8.0E-2
13	2.7185059	8.2E-5	2.7182617	-7.4E-6	2.0000000	-2.6E-1
14	2.7182617	-7.4E-6	2.7182617	-7.4E-6	0.	-1.0E+0
15	2.7177734	-1.9E-4	2.7187500	1.7E-4	-32.000000	-1.3E+1
16	2.7187500	1.7E-4	2.7187500	1.7E-4	0.	-1.0E+0
17	2.7187500	1.7E-4	2.7187500	1.7E-4	0.	-1.0E+0
18	2.7187500	1.7E-4	2.7187500	1.7E-4	0.	-1.0E+0

例 6

	前半周期	後半周期	一周期
真値	2.912411653	0.7151870750	3.627598728
4分割	2.617993879	1.047197551	3.665191429
	(-0.294417774)	(0.332010476)	(0.037592701)
	2.792526804	0.6981317007	3.490658504
	(-0.119884849)	(-0.0170553743)	(-0.136940224)
8分割	2.841292404	0.7864991142	3.627791519
	(-0.071119249)	(0.0713120392)	(0.000192791)
	2.915725245	0.6995996350	3.615324883
	(0.003313592)	(-0.0155874400)	(-0.012273845)
12分割	2.881493711	0.7461060118	3.627599724
	(-0.030917942)	(0.0309189368)	(0.000000996)
	2.914490922	0.7122137621	3.626704684
	(0.002079269)	(-0.0029733129)	(-0.000894044)
24分割	2.904769611	0.7228291182	3.627598734
	(-0.007642042)	(0.0076420432)	(0.000000006)
	2.912528245	0.7150701537	3.627598398
	(0.000116592)	(-0.0001169213)	(-0.000000330)

前半周期 $\int_0^{\pi} \frac{1+\sin x}{2+\cos x}\,dx$

後半周期 $\int_{\pi}^{2\pi} \frac{1+\sin x}{2+\cos x}\,dx$

一周期 $\int_0^{2\pi} \frac{1+\sin x}{2+\cos x}\,dx$

上段： 台形公式
下段： シンプソン公式
( ) 内は誤差

例 7

a b c

$f_2(X)=1.0000X^2-314.19X+9.8696=0$

$b^2=98715.$ $\sqrt{r}=314.13$

$4ac=39.478$ $-b=314.19$

$r=b^2-4ac$ $-b+\sqrt{r}=628.32$

$=98676.$ $-b-\sqrt{r}=0.060000$

$X_1=314.16$ $X_2=0.030000$

$\bar{a}=1.000014142$ $\bar{b}^2=98716.44446$

$\bar{b}=-314.1917320$ $4\bar{a}\bar{c}=39.47904774$

$\bar{c}=9.869622360$ $\bar{r}=98676.96541$

$\sqrt{\bar{r}}=314.1288994$ $-b+\sqrt{\bar{r}}=628.3206314$

$-\bar{b}=314.1917320$ $-b-\sqrt{\bar{r}}=0.06283260000$

$\bar{X}_1$ 314.1558729 $\bar{X}_2=0.03141585572$

—： 任意の数値を加えた桁に影響される桁

例8　2倍桁の3元連立1次方程式（ガウスの消去法）

ε_{11}	$a_{12}^{(0)}$	$a_{13}^{(0)}$	$c_1^{(0)}$
4.949E-4	-3.000E+0	4.000E+0	-2.000E+0
$a_{21}^{(0)}$	$a_{22}^{(0)}$	$a_{23}^{(0)}$	$c_2^{(0)}$
-4.000E+0	3.000E+0	2.000E+0	4.000E+0
$a_{31}^{(0)}$	$a_{32}^{(0)}$	$a_{33}^{(0)}$	$c_3^{(0)}$
3.000E+0	-4.000E+0	2.000E+0	-3.000E+0

$a_{11}^{(1)} = \varepsilon_{11} / \varepsilon_{11}$	$a_{12}^{(1)} = a_{12}^{(0)} / \varepsilon_{11}$	$a_{13}^{(1)} = a_{13}^{(0)} / \varepsilon_{11}$	$c_1^{(1)} = c_1^{(0)} / \varepsilon_{11}$
1.0	-6.062E+3	8.082E+3	-4.041E+3

$\bar{a}_{11}^{(1)}$	$\bar{a}_{12}^{(1)}$	$\bar{a}_{13}^{(1)}$	$\bar{c}_1^{(1)}$
1.0	-6.0618586E+3	8.0821732E+3	-4.0407764E+3
$\bar{a}_{21}^{(0)}$	$\bar{a}_{21}^{(0)} \cdot \bar{a}_{12}^{(1)}$	$\bar{a}_{21}^{(0)} \cdot \bar{a}_{13}^{(1)}$	$\bar{a}_{21}^{(0)} \cdot \bar{c}_1^{(1)}$
-3.9996859E+0	2.4245530E+4	-3.2326154E+4	1.6161836E+4
$\bar{a}_{31}^{(0)}$	$\bar{a}_{31}^{(0)} \cdot \bar{a}_{12}^{(1)}$	$\bar{a}_{31}^{(0)} \cdot \bar{a}_{13}^{(1)}$	$\bar{a}_{31}^{(0)} \cdot \bar{c}_1^{(1)}$
3.0002718E+0	-1.8187223E+4	2.4248716E+4	-1.2123427E+4

$\bar{a}_{22}^{(1)} = \bar{a}_{22}^{(0)} - \bar{a}_{21}^{(0)} \cdot \bar{a}_{12}^{(1)}$	$\bar{a}_{23}^{(1)} = \bar{a}_{23}^{(0)} - \bar{a}_{21}^{(0)} \cdot \bar{a}_{13}^{(1)}$	$\bar{c}_2^{(1)} = \bar{c}_2^{(0)} - \bar{a}_{21}^{(0)} \cdot \bar{c}_1^{(1)}$
-2.4242530E+4	3.2328154E+4	-1.6157836E+4
$\bar{a}_{32}^{(1)} = \bar{a}_{32}^{(0)} - \bar{a}_{31}^{(0)} \cdot \bar{a}_{12}^{(1)}$	$\bar{a}_{33}^{(1)} = \bar{a}_{33}^{(0)} - \bar{a}_{31}^{(0)} \cdot \bar{a}_{13}^{(1)}$	$\bar{c}_3^{(1)} = \bar{c}_3^{(0)} - \bar{a}_{31}^{(0)} \cdot \bar{c}_1^{(1)}$
1.8183223E+4	-2.4246716E+4	1.2120427E+4

$a_{ij}^{(k)}$ ： 単精度

$\bar{a}_{ij}^{(k)}$ ： 2倍精度

—— ： 2倍精度のときの下位4桁

$\bar{a}_{22}^{(2)}=\bar{a}_{22}^{(1)}/\bar{a}_{22}^{(1)}$	$\bar{a}_{23}^{(2)}=\bar{a}_{23}^{(1)}/\bar{a}_{22}^{(1)}$	$\bar{c}_{2}^{(2)}=\bar{c}_{2}^{(1)}/\bar{a}_{22}^{(1)}$
1.0	-1.3335305E+0	6.6650783E-1
$\bar{a}_{32}^{(1)}$	$\bar{a}_{32}^{(1)}\cdot\bar{a}_{23}^{(2)}$	$\bar{a}_{32}^{(1)}\cdot\bar{c}_{2}^{(2)}$
1.8183223E+4	-2.4247882E+4	1.2119261E+4

$\bar{a}_{33}^{(2)}=\bar{a}_{33}^{(1)}-\bar{a}_{32}^{(1)}\cdot\bar{a}_{23}^{(2)}$	$\bar{c}_{3}^{(2)}=\bar{c}_{3}^{(1)}-\bar{a}_{32}^{(1)}\cdot\bar{c}_{2}^{(2)}$
1.1660000E+0	1.1660000E+0

$\bar{a}_{13}^{(1)}\cdot\bar{X}_3$
8.0839779E+3
$\bar{a}_{12}^{(1)}\ \bar{X}_2$
-1.2125755E+4

$\bar{a}_{12}^{(1)}\cdot\bar{X}_2+\bar{a}_{13}^{(1)}\cdot\bar{X}_3$	$\bar{a}_{23}^{(2)}\cdot\bar{X}_3$	
-4.0417771E+3	-1.3338283E+0	
$\bar{X}_1$	$\bar{X}_2=\bar{c}_{2}^{(2)}-a_{23}^{(2)}\cdot\bar{X}_3$	$X_3=c_3^{(2)}/a_{33}^{(2)}$
1.0007000E+0	2.0003361E+0	1.000E+0

X_1	X_2	$\bar{X}_3$
1.001E+0	2.000E+0	1.0002233E+0

真の解は $X_1=1.0$, $X_2=2.0$, $X_3=1.0$

例 9

$f_3(X)=(X-\pi\times10^4)(X-\pi\times10^2)(X-\pi)=0$

$\bar{f}_3(X)=1.000X^3-3.173E+4X^2+9.969E+6X-3.101E+7=0$

X_i	3.141267545E+4	3.14183E+2
max i	2	2 or 1
$\lvert a_{\max i} X_i^{\max i} \rvert$	3.13E+13	3.13E+9
$\lvert a_0 \rvert \cdot 10^{-4}$	3.10E+3	3.10E+3
計算桁数	13-3=10	9-3=6
$b_2 \cdot X_i$	* ********* 3.141267545E+4	* ***** 3.14183E+2
a_2	* ********* -3.173000000E+4	* ***** -3.17300E+4
b_1	* ********'' 3.173245500E+2	* ***** 3.14158E+4
$b_1 \cdot X_i$	* ********'' -9.968013101E+6	* ***** -9.87031E+6
a_1	* ********* 9.969000000E+6	* ***** 9.96900E+6
b_0	* ***'''''' 9.868990000E+2	* ***'' 9.86900E+4
$b_0 \cdot X_i$	* ***'''''' 3.100113799E+7	* ***'' 3.10067E+7
a_0	* ********* -3.101000000E+7	* ***** -3.10100E+7
b_{-1}	' ''''''''' -8.862010000E+3	' ''''' -3.30000E+3
b_2	1.000	1.000
b_1	-3.173E+2	-3.142E+4
b_0	9.869E+2	9.869E+4
b_1^2	1.007E+5	9.872E+8
$4b_2 \cdot b_0$	3.948E+3	3.948E+5
$b_1^2-4b_2 \cdot b_0$	9.675E+4	9.868E+8
$\sqrt{b_1^2-4b_2 \cdot b_0}$	3.110E+2	3.141E+4
$-b_1+\sqrt{b_1^2-4b_2 \cdot b_0}$	6.283E+2	6.283E+4
$2b_0$	1.974E+3	1.974E+5
X_2 or X_1	3.142E+2	3.142E+4
X_3	3.141	3.142

* : 丸めの誤差の入っていない桁

' : 丸めの誤差の入っている桁

＿ : 任意の数値を加えた桁に影響される桁

例 10

$$(X-3.142)(X^2-9.870)=(5.507X-17.30)^2$$

$\sum_{i=0}^{3} a_i X^i=0$　　$a_3=1.000$　　$a_2=-33.47$

$a_1=180.6$　　$a_0=-268.3$

$X_1=27.19$　　$X_2=3.139+0.1225i$

$X_3=3.139-0.1225i$

$X_1+X_2+X_3=33.47$

$X_1X_2X_3=268.3$

$X_1X_2+X_2X_3+X_1X_3=180.6$

X	27.19	$3.139 \pm 0.1225i$
a_3X^3	20100.	$30.79 \pm 3.619i$
a_2X^2	-24740.	$-329.3 \mp 25.74i$
a_1X	4911.	$566.9 \pm 22.12i$
a_0	-268.3	-268.3
$f(X)$	2.700	$0.09000 \mp 0.001000i$

$$(X-3.1421732)(X^2-9.8702449)=(5.5072236X-17.301414)^2$$

$\sum_{i=0}^{3} \bar{a}_i X^i=0$　　$\bar{a}_3=1.0000000$　　$\bar{a}_2=-33.471685$

$\bar{a}_1=180.69527$　　$\bar{a}_0=-268.32491$

$X_1=27.188696$　　$X_2=3.1419946$　　$X_3=3.1409946$

$X_1+X_2+X_3=33.471685$

$X_1X_2+X_2X_3+X_1X_3=180.69527$

$X_1X_2X_3=268.32492$

～～～：計算に使用されていない桁

X	3.1419946	3.1409946
$\bar{a}_3 X^3$	31.018179	30.988572
$\bar{a}_2 X^2$	-330.43683	-330.22653
$\bar{a}_1 X$	567.74356	567.56287
$\bar{a}_0$	-268.32491	-268.32491
$f(X)$	-0.0000010000000	0.0000020000000

— : 任意の数値を加えた桁に影響される桁

例 11

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & 2 & 3 & 4 \ \cdots\cdots & 12 \end{pmatrix}$$

$a_{ij}=\min(i,j)$

$(i,j=1,2,\ \cdots\cdots\ ,12)$

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -r_0 \\ 1 & 0 & 0 & \cdots & 0 & -r_1 \\ 0 & 1 & 0 & \cdots & 0 & -r_2 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 0 & -r_{10} \\ 0 & 0 & 0 & \cdots & 1 & -r_{11} \end{pmatrix}$$

$a_{i\,i-1}=1 \quad (i=2,3,\ \cdots\cdots\ ,12)$

$a_{i\,12}=-r_{i-1} \quad (i=1,2,3,\ \cdots\cdots\ ,12)$

$a_{ij}=0 \quad i\neq j+1$

$\begin{pmatrix} i=1,2,\ \cdots\cdots\ ,12 \\ j=1,2,\ \cdots\cdots\ ,11 \end{pmatrix}$

$$\lambda^{12}+\sum_{i=0}^{11} r_i \lambda^i = 0$$

r_{11} = -77.99999997209099765536	λ_1 =63.4091389125.....
r_{10} = 1001.000000015736826733	λ_2 = 7.1201221912.....
r_9 = -5004.999999555429417040	λ_3 = 2.6180339580.....
r_8 = 12870.000000694779977791	λ_4 = 1.379021273.....
r_7 =-19448.000001027806374 43	λ_5 = 0.870745072.....
r_6 = 18564.000002564831113 60	λ_6 = 0.615295469.....
r_5 =-11628.000002137646246 98	λ_7 = 0.47045769.....
r_4 = 4845.000001219383626031	λ_8 = 0.38197018.....
r_3 = -1330.000000384419967809	λ_9 = 0.32555041.....
r_2 = 231.000000079726977 4931	λ_{10}= 0.2891984 9.....
r_1 = -23.000000008478378847 86	λ_{11}= 0.2664741 3.....
r_0 = 1.000000000397431910934	λ_{12}= 0.2539921 7.....

r_{11} = -78.000000000248537533 39	λ_1 =63.4091389487 13.....
r_{10} = 1001.000000000125698504	λ_2 = 7.1201221745 467.....
r_9 = -5004.999999963405295403	λ_3 = 2.6180339887 398.....
r_8 = 12869.999999780009326 95	λ_4 = 1.3790211868 893.....
r_7 =-19447.999999457183647 96	λ_5 = 0.8707453295 243.....
r_6 = 18563.999999275814241 94	λ_6 = 0.6152947365 865.....
r_5 =-11627.999999417715870 88	λ_7 = 0.4704595974 377.....
r_4 = 4844.999999703969466984	λ_8 = 0.3819660112 68.....
r_3 = -1329.999999903884624250	λ_9 = 0.3255575444 20.....
r_2 = 230.999999980627709 4679	λ_{10}= 0.2891897470 40.....
r_1 = -22.999999997787223890 79	λ_{11}= 0.2664809571 40.....
r_0 = 0.999999999890384120928 9	λ_{12}= 0.2539897779 40.....

— : 無効桁

例 12

$$A_T=\begin{pmatrix} 0.876929..... & 0.702590..... & 0.487499..... \\ 0.702590..... & 0.579350..... & 0.413240..... \\ 0.487499..... & 0.413240..... & 0.303266..... \end{pmatrix}$$

$\lambda_{1T}=\sqrt{3}$　　$\lambda_{2T}=e\cdot 10^{-2}$　　$\lambda_{3T}=\pi\cdot 10^{-4}$

$$V_{1T}=\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{3} \\ 1/\sqrt{6} \end{pmatrix}\qquad V_{2T}=\begin{pmatrix} \sqrt{2/5} \\ -1/\sqrt{15} \\ -\sqrt{8/15} \end{pmatrix}\qquad V_{3T}=\begin{pmatrix} -1/\sqrt{10} \\ \sqrt{6/10} \\ -\sqrt{3/10} \end{pmatrix}$$

$$A_0=\begin{pmatrix} 0.8769 & 0.7026 & 0.4875 \\ 0.7026 & 0.5794 & 0.4132 \\ 0.4875 & 0.4132 & 0.3033 \end{pmatrix}$$

$$V_{iT}\,V_{jT}^{T}=\begin{cases} 1.0 & (i=j) \\ 0.0 & (i\neq j) \end{cases}\qquad A_T=\sum_{i=1}^{3}\lambda_{iT}V_{iT}\,V_{iT}^{T}$$

$$\bar{V}_{01}^{*}=\begin{pmatrix} 1.0 \\ 1.0 \\ 1.0 \end{pmatrix}\longrightarrow\bar{V}_{11}=\begin{pmatrix} 0.7050085187 \\ 0.5782004461 \\ 0.4106668147 \end{pmatrix}\dashrightarrow\bar{V}_{q1}=\begin{pmatrix} 0.7070954160 \\ 0.5773632883 \\ 0.4082495625 \end{pmatrix}$$

$\bar{A}_1=A_0-\bar{\lambda}_1\bar{V}_{q1}\,\bar{V}_{q1}^{T}$　　$\bar{\lambda}_1=1.732082329$

$$\bar{A}_1=\begin{pmatrix} 0.01089716280 & -0.004513274900 & -0.01249121900 \\ -0.004513274900 & 0.002028017000 & 0.004948967400 \\ -0.01249121900 & 0.004948967400 & 0.01463597870 \end{pmatrix}$$

— : 無効桁

$$\bar{V}_{02}^{*}=\begin{pmatrix}1.0\\1.0\\1.0\end{pmatrix}\longrightarrow\bar{V}_{12}=\begin{pmatrix}-0.6309633009\\0.2545318492\\0.7328702822\end{pmatrix}\cdots\rightarrow\bar{V}_{52}=\begin{pmatrix}-0.6320796784\\0.2572527471\\0.7309557469\end{pmatrix}$$

$$\bar{A}_2=\bar{A}_1-\bar{\lambda}_2\bar{V}_{52}\bar{V}_{52}^{T}\qquad\bar{\lambda}_2=0.02717925671$$

$$\bar{A}_2=\begin{pmatrix}3.837788000E\text{-}5 & -9.381269000E\text{-}5 & 6.620297000E\text{-}5\\ -9.381269000E\text{-}5 & 2.293216250E\text{-}4 & -1.618301940E\text{-}4\\ 6.620297000E\text{-}5 & -1.618301940E\text{-}4 & 1.142023000E\text{-}4\end{pmatrix}$$

$$\bar{V}_{03}^{*}=\begin{pmatrix}1.0\\1.0\\1.0\end{pmatrix}\longrightarrow\bar{V}_{13}=\begin{pmatrix}0.3170130201\\-0.7748939287\\0.5468474596\end{pmatrix}\cdots\rightarrow\bar{V}_{33}=\begin{pmatrix}0.3170034505\\-0.7749019029\\0.5468417079\end{pmatrix}$$

$$\bar{A}_3=\bar{A}_2-\bar{\lambda}_3\bar{V}_{33}\bar{V}_{33}^{T}\qquad\bar{\lambda}_3=0.0003819015137$$

$$\bar{A}_3=\begin{pmatrix}1.433400000E\text{-}10 & 1.095900000E\text{-}10 & 7.210000000E\text{-}11\\ 1.095900000E\text{-}10 & 9.300000000E\text{-}11 & 6.830000000E\text{-}11\\ 7.210000000E\text{-}11 & 6.830000000E\text{-}11 & 5.490000000E\text{-}11\end{pmatrix}$$

＊の付いたベクトルは正規化されていない。